



Силабус освітнього компонента Програма навчальної дисципліни



Математична статистика

Шифр та назва спеціальності

122 – Комп'ютерні науки

Інститут

ННІ Комп'ютерного моделювання, прикладної фізики та математики

Освітня програма

Комп'ютерні науки. Моделювання, проектування та комп'ютерна графіка

Кафедра

Математичне моделювання та інтелектуальні обчислення в інженерії (161))

Рівень освіти

Бакалавр

Тип дисципліни

Спеціальна(фахова), Обов'язкова

Семестр

5

Мова викладання

Українська, Англійська

Викладачі, розробники



Ларін Олексій Олександрович
(відповідальний лектор)

Oleksiy.Larin@khpi.edu.ua

Доктор технічних наук, професор

Фахівець в галузі обчислювальної математики, ймовірного моделювання та прогнозуванні надійності. Основний фокус наукових робіт присвячено розробці моделей, методів, підходів та алгоритмів комп'ютерного моделювання та статистичного аналізу інженерних систем, зокрема із випадковими параметрами. Автор понад 150 наукових та методичних праць.

[Детальніше про викладача на сайті кафедри](#)



Потопальська Ксенія Євгенівна
(асистент з лабораторних робіт)

Kseniia.Potopalska@khpi.edu.ua

Кандидат технічних наук

[Детальніше про викладача на сайті кафедри](#)



Некрасова Марія Володимирівна
(асистент з лабораторних робіт)

marii.nekrasova@khpi.edu.ua

Кандидат технічних наук

[Детальніше про викладача на сайті кафедри](#)

Загальна інформація

Анотація

Курс статистики спрямований на формування теоретичних знань та практичних навичок у студентів для вирішення прикладних проблем у галузі математичної статистики. Здобувачі освіти розвивають абстрактне мислення, аналітичні здібності та програмувальні навички, використовуючи ймовірісно-статистичні методи аналізу даних, прогнозування та прийняття обґрунтованих рішень. Курс включає лекції, комп'ютерний лабораторний практикум та командну проектну роботу.

Основні теми включають вивчення описової статистики, основ статистичної діагностики та статистичного аналізу даних. Командний проєкт дозволяє студентам застосовувати отримані знання на практиці та демонструвати їхні навички вирішення прикладних завдань.

Курс враховує сучасні інформаційні технології та надає студентам можливість набути досвід використання Python для вирішення задач математичної статистики. Підсумковий контроль – іспит.

Мета та цілі дисципліни

Мета курсу математичної статистики полягає в формуванні у майбутніх фахівців теоретичних знань та практичних навичок зі застосування ймовірісно-статистичних методів для розв'язання прикладних проблем в галузі моделювання, аналізу даних, прогнозування та прийняття рішень, зокрема зі застосуванням сучасних інформаційних інструментів та під час роботи в команді.

При цьому ставляться наступні цілі:

1. Надання здобувачам освіти знань та уявлень щодо використання математичних концепцій та методів для статистичного аналізу даних, розробки прогнозів та прийняття статистично обґрунтованих рішень у різних галузях.
2. Розвиток абстрактного мислення та аналітичних навичок, дозволяючи студентам впроваджувати ймовірісно-статистичні методи в практику.
3. Надання практичного досвіду реалізації алгоритмів статистичного аналізу та покращення навичок програмування.
4. Покращити здатність самостійно формалізувати та розв'язувати завдання, використовуючи ймовірісно-статистичні методи, а також ефективно співпрацювати в команді та взаємодіяти з іншими фахівцями в обраній галузі.

Формат занять

Лекції, лабораторні роботи, самостійна робота. Підсумковий контроль – іспит.

Компетентності

ЗК2: Здатність застосовувати знання у практичних ситуаціях

ЗК9: Здатність працювати в команді.

ЗК11: Здатність приймати обґрунтовані рішення.

СК7: Здатність застосовувати теоретичні та практичні основи методології та технології моделювання для дослідження характеристик і поведінки складних об'єктів і систем, проводити обчислювальні експерименти з обробкою й аналізом результатів

Результати навчання

ПР1: Застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів вилучення, аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук.

ПР3: Використовувати знання закономірностей випадкових явищ, їх властивостей та операцій над ними, моделей випадкових процесів та сучасних програмних середовищ для розв'язування задач статистичної обробки даних і побудови прогнозних моделей.

ПР8: Використовувати методологію системного аналізу об'єктів, процесів і систем для задач аналізу, прогнозування, управління та проектування динамічних процесів в макроекономічних, технічних, технологічних і фінансових об'єктах

Обсяг дисципліни

Загальний обсяг дисципліни 120 год. (4 кредита ECTS): лекції – 32 год., лабораторні роботи – 16 год., самостійна робота – 72 год.

Передумови вивчення дисципліни (пререквізити)

Основним реквізитом до цього курсу є вивчення дисципліни "Теорія ймовірностей".

Важливими є також знання основ математичного аналізу в частині диференційного та інтегрального числень та основ лінійної алгебри.

Курс передбачає комп'ютерний лабораторний практикум тому навички програмування («Об'єктно - орієнтоване програмування та проектування») є також необхідними для його успішного проходження (початковий рівень володіння Python вітається проте не є обов'язковим)

Особливості дисципліни, методи та технології навчання

Методи викладання та вивчення курсу включають традиційні лекції для теоретичних основ, а також для закріплення практичних навичок запропоновано комп'ютерний лабораторний практикум, в рамках якого студенти виконують завдання статистичного аналізу та діагностики з використанням мови Python у середовищі PyCharm. Виконання лабораторних завдань передбачено переважно в рамках самостійної роботи в той час як на лабораторних заняттях відбувається консультування та захист виконаних робіт.

В рамках курсу студенти також здійснюють обов'язкове розрахункове завдання, яке подається у формі командного проекту. Виконання проектної роботи та її захист є обов'язковим елементом навчальної дисципліни, результати якої входять до формування підсумкової оцінки.

Програма навчальної дисципліни

Теми лекційних занять

Тема 1. Основні поняття математичної статистики.

Основні визначення математичної статистики. Постановка задач описової статистики. Аналітична статистика: діагностика, прогнозування, прийняття рішень.

Тема 2. Статистичні гіпотези.

Загальні поняття щодо 0 та альтернативної гіпотез. Помилки першого та другого роду. Прийняття статистичних рішень: рівень значущості та р-значення.

Тема 3. Задача статистичної діагностики.

Розглядається задача прийняття статистичних рішень на основі застосування узагальненої формули Байеса.

Тема 4. Статистичні характеристики вибірки та їх точкові оцінки

Поняття генеральних значень та вибіркової оцінки для числових характеристик розподілів вибірки. Вимоги до точкових оцінок. Приклади задоволення вимог для оцінок. Визначення спроможності, незміщеності та ефективності статистичних оцінок.

Тема 5. Точкові оцінки для математичного сподівання та дисперсії

Статистичні оцінки для математичного сподівання та дисперсії. Теоретичні оцінки для їх спроможності та незміщеності

Тема 6. Довірчі оцінки та інтервальне оцінювання

Принципи побудови довірчих інтервалів. Квантілі та квартилі розподілу. Односторонні та двосторонні оцінки.

Тема 7. Інтервальна оцінка математичного сподівання та дисперсії

Побудова довірчих інтервалів для математичного сподівання. Критерій ст'юдента. Критерій хі-квадрат Пірсона

Тема 8. Порівняння вибірок

Постановка задачі. Порівняння 2 дисперсій (F-критерій Фішера). Порівняння математичних сподівань. Порівняння кількох вибірок (визначення понять щодо критерію Бартлета та Кочрена)

Тема 9. Статистичні розподіли вибірки. Підбір розподілу

Алгоритм побудови гістограми. Принципи підбору розподілу. Метод моментів та використання принципу максимуму правдоподібності

Тема 10. Непараметричні статистичні гіпотези.

Перевірка правильності вибору розподілу. Критерії узгодженості Колмогорова та Пірсона.

Тема 11. Елементи дисперсійного аналізу

Поняття дисперсійного аналізу. Однофакторний дисперсійного аналіз. Особливості дво та багатфакторного дисперсійного аналізу.

Тема 12. Лінійний регресійний аналіз.

Поняття регресійного аналізу. Теоретичні основи методу найменших квадратів, сформульовані відносно принципу максимальної правдоподібності..

Тема 13. Узагальнений лінійний регресійний аналіз.

Узагальнення методу лінійного регресійного аналізу на застосування різних систем ортнормованих базисних функцій. Поліноміальна та тригонометрична регресія.

Тема 14. Статистичний аналіз регресійної моделі

Інтервальні оцінки для параметрів регресії, довірчі інтервали та оцінки похибки.

Тема 15. Елементи кореляційного аналізу

Поняття стохастичного зв'язку та кореляції. Точкові оцінки вибіркового коефіцієнту кореляції. Кореляційна матриця

Тема 16. Елементи теорії прийняття статистичних рішень

Визначення статистичного ризику та його використання для прийняття рішень. Підхід мінімізації емпіричного ризику та мінімаксий підхід.

Теми практичних занять

Не передбачено навчальним планом.

Теми лабораторних робіт

Тема 1. Знайомство з мовою Python та середовищем PyCharm

Побудова апроксимацій. Візуалізації даних.

Тема 2. Вирішення задачі діагностики методом Байеса.

Пропонується елементарне написання програмного коду для задачі, яка розібрана на практичному занятті (фактично "ручна" реалізації формул)

Тема 3. Вирішення задачі медичної діагностики методом Байеса.

Для заданого датасету щодо даних по симптомам пацієнтів з кількома діагнозами слід реалізувати програмно систему підтримки діагностичних рішень

Тема 3а. Створення графічного інтерфейсу для системи, яка проводить автоматизовану медичну діагностику (необов'язкова робота).

Для попередньої лабораторної роботи слід створити програмне рішення з орієнтацією на користувача щодо інтерфейсу взаємодії та візуалізації результатів.

Тема 4. Підбір розподілу до вибірки. Перевірка непараметричних гіпотез

Для заданого датасету (вибірки) необхідно обчислити точкові оцінки, побудувати гістограму, підібрати розподіл та перевірити критерії узгодженості

Тема 5. Регресійний аналіз

Для заданого датасету (вибірки) побудувати просту лінійну регресію та отримати довірчі інтервали до неї (передбачається "ручна реалізація"). З використанням бібліотек побудувати поліноміальну, тригонометричну та багатовимірну регресію.

Тема 6. Кореляційний аналіз.

Для заданого датасету побудувати та візуалізувати кореляційну матрицю

Тема 7. Прийняття статистичних рішень

Для наявного датасету визначити граничних значень для прийняття рішення щодо діагностичного параметру в оцінці вірусного захворювання пацієнтів.

Захист командних проєктів

Самостійна робота

Опрацювання лекційного матеріалу.

Виконання лабораторних робіт, зокрема додаткових індивідуальних завдань та запитань до них.

Виконання командної проєктної роботи.

Вимоги щодо форми проєктної роботи та процедури її виконання:

1) Проєктна робота виконується командою з 2–4 студентів.

- 2) Початок виконання роботи на 12-тому тижні і термін її виконання 3 тижня.
- 3) Форма звітності: публічний захист із презентацією результатів виконання проекту та демонстрацією програмних засобів, якими було досягнуто ці результати.
- 4) Атестацію проектних робіт виконує комісія, що складається з лектора дисципліни та викладачів, які здійснюють проведення практичних/лабораторних робіт.
- 5) Студенти можуть вільно обирати програмно-технічні засоби для виконання даної роботи.

Основні вимоги до проектної роботи по суті:

Робота передбачає здійснення студентами статистичного аналізу даних. Команда має обрати масив даних, з певної предметної області з об'ємом порядку 1000 значень або більше. Предметна область (тематика) проекту може обиратись командою студентів довільно з використанням відкритих джерел даних (це можуть бути дані щодо екології та її зміни, дані спортивних змагань, відкритих даних по фінансовим, ринковим питанням, медична статистика, інфраструктурне забезпечення, безпека тощо). Наведений перелік є дуже умовним і не має на меті обмежити вибір студентів, але тема має бути погоджена з лектором перед початком виконання проекту.

За результатом виконання проекту команда може отримати до 20 балів до підсумкової атестації.

Оцінюється проект в цілому, тобто уся команда, але оцінка відповідей на запитання є персональною для кожного студента члена команди.

Для самостійного опрацювання студентам рекомендується курси неформальної освіти на платформі Coursera :

Inferential Statistical Analysis with Python (University of Michigan)

Fitting Statistical Models to Data with Python (University of Michigan)

<https://www.coursera.org/learn/inferential-statistical-analysis-python?specialization=statistics-with-python>

Література та навчальні матеріали

1. Сеньо П. С. Теорія ймовірностей та математична статистика: підручник. – К.: Знання, 2007, 556с.
2. Турчин В.М. Теорія ймовірностей і математична статистика. Основні поняття, приклади, задачі. – Д.: ІМАпрес, 2014.
3. Драгомирецька Х.Т. та ін Теорія ймовірностей та математична статистика. – Львів: Львівська політехніка, 2012, 389 с.
4. Ron Kenett, Shelemyahu Zacks, Peter Gedeck Modern Statistics. A Computer-Based Approach with Python, Springer Birkhäuser, 2022, Switzerland, 438 p.
5. Щоголев С. А. Основи теорії ймовірностей та математичної статистики: навчально-методичний посібник. – Одеса : «Одеський національний університет ім. І. І. Мечникова», 2015. – 206 с.
6. А. В. Скороход Теорія ймовірностей та математична статистика К. : ТВіМС, 1995
7. Методичні вказівки до практичних занять з курсу "Теорія ймовірностей, ймовірнісні процеси і математична статистика" для студентів спеціальності "Інформаційні технології проектування" / уклад.: М. В. Матюшенко, Г. В. Федченко, І. Б. Шеліхова. — Харків: НТУ "ХПІ", 2015. — 35с
8. Методичні вказівки до виконання індивідуальних завдань комп'ютерного практикуму на Python з курсу «Математична статистика» для студентів технічних дисциплін частина 1 / уклад. О.О. Ларін, О. І. Суханова – Харків : НТУ «ХПІ». – 2020, 48 с.
9. Карташов М.В. Імовірність, процеси, статистика. К.: Київський університет, 2008. – 494 с

Система оцінювання

Критерії оцінювання успішності студента та розподіл балів

Теоретична частина: 20 балів

у формі усного екзамену (2 питання у білеті)

Практична лабораторна частина курсу: 60 балів

8 лабораторних робіт, що в залежності від

складності оцінюються від 5 до 10 балів

Оцінка є рекомендацією викладача, що проводить лабораторні роботи.

Виконання лабораторних робіт є допуском до екзамену.

Командна проектна робота: 20 балів

Повнота розкриття проблеми 3..5 балів

Математична складність 3..5

Програмно-алгоритмічні рішення 0..2

Якість презентації 1..3

Відповіді на питання комісії 0..5

(індивідуально до кожного члена команди)

Виконання рекомендованих курсів неформальної

освіти (див розділ Самостійна робота) дозволяє

студенту після співбесіди на результатів навчання

по окремим лабораторним роботам та\або

проєкту в обсязі до 20 балів.

Шкала оцінювання

Сума балів	Національна оцінка	ECTS
90–100	Відмінно	A
82–89	Добре	B
75–81	Добре	C
64–74	Задовільно	D
60–63	Задовільно	E
35–59	Незадовільно (потрібне додаткове вивчення)	FX
1–34	Незадовільно (потрібне повторне вивчення)	F

Норми академічної етики і політика курсу

Студент повинен дотримуватися «Кодексу етики академічних взаємовідносин та доброчесності НТУ «ХПІ»: виявляти дисциплінованість, вихованість, доброзичливість, чесність, відповідальність.

Конфліктні ситуації повинні відкрито обговорюватися в навчальних групах з викладачем, а при неможливості вирішення конфлікту – доводитися до відома співробітників дирекції інституту.

Нормативно-правове забезпечення впровадження принципів академічної доброчесності НТУ

«ХПІ» розміщено на сайті: <http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/>

Погодження

Силабус погоджено

28.08.2023

Завідувач кафедри
Олексій ВОДКА

28.08.2023

Гарант ОП
Оксана ТАТАРІНОВА